Hostile Scaffolding

Ryan Timms & David Spurrett
(UKZN)
ryangarytimms@gmail.com / spurrett@ukzn.ac.za

Abstract

Discussion of cognitive scaffolding is dominated by attention to ways that external structure can support cognitive activity or augment an agent's cognitive capacities. We call instances where the interests of the user are served *benign* and argue for the possibility of *hostile* scaffolding. This is scaffolding which depends on the same capacities of an agent to rely on external structure, but that undermines or exploits that agent while serving the interests of another. We offer one defence of hostile scaffolding by developing an account of a neglected complementarity between extended phenotype thinking and extended functionalism. We support this with a second defence, an account of design features of electronic gambling machines and casino management systems that show how they exemplify hostile scaffolding.

Keywords:

Scaffolding, hostility, extended functionalism, extended phenotypes

1. Introduction

The claim that cognition can be *scaffolded* occurs frequently in discussion of whether minds or cognition are extended.[1] The general idea is that things in the environment, including other agents, can support, simplify or otherwise beneficially transform cognitive processes just as physical scaffolding can beneficially transform the demands and risks of construction and maintenance. The idea of cognitive scaffolding goes back at least to Vygotsky, who characterized a 'zone of proximal development' indicating, among other things, a learning stage where a task could be performed successfully only with a supportive guide (Vygotsky 1978).[2] In contemporary cognitive science 'scaffolding' is used to embrace non-temporary cognitive infrastructure, and resources outside explicitly developmental contexts, including everyday supports like scheduling tools and conventionally labelled public spaces. Humans are distinguished by the extent of their reliance on scaffolding, and in the great variety of forms of scaffolding they use (Clark 1997, Sterelny 2012).

We argue here for the theoretical coherence, actual existence, and importance of *hostile* cognitive scaffolding. (From here on we'll mostly drop the 'cognitive' qualifier.) Most discussion of scaffolding focuses on how it can be supportive or helpful by reducing cognitive demands, improving reliability, facilitating co-operation, and making possible tasks that would be beyond an unaided individual. We don't dispute that this *benign* scaffolding is genuine and important. But agents that are capable of benefitting from benign scaffolding may for that reason be vulnerable to external structure that interfaces with their capacities for distributing cognitive labour in ways that harm them. When this vulnerability is exploited in ways that undermine their own interests *and* serve those of other agents, we call the scaffolding *hostile*. We mean hostile in the sense of Sterelny (2003), whose emphasis on informational hostility is an important corrective to optimistic slogans about the world serving as 'its own best representation' (Brooks 1991).

Our defence of hostile scaffolding has three parts. One develops an account of what hostile scaffolding is, so that we have something definite to aim for. This involves clarifying

1 The claim that cognition can be extended, including being scaffolded is weaker than and different to the claim that *minds* are extended (Sterelny 2010). Our interest here is in cognitive scaffolding, not the boundaries of minds.

2 The term 'scaffolding' (in English) to refer to cognitive developmental support was introduced in Wood, Bruner and Ross (1976), and first explicitly linked to Vygotsky in Cazden (1981).

what scaffolding is, distinguishing benign from hostile scaffolding, and introducing a distinction between shallow and deep scaffolding (section 2). The second develops an account of complementarities between arguments for extended cognition, and the defence of the idea of extended phenotypes which together predict hostile scaffolding (section 3). The third is a case study, arguing that contemporary machine gambling devices and casino management systems are real instances of hostile scaffolding (section 4). Sterelny (2010) has argued, without using the label, that hostile scaffolding is unlikely to be a big deal. We respond to his reasons for skepticism late in section (4), and conclude in Section (5).

Our claims are not entirely unprecedented. Slaby (e.g. 2016) has argued that affective scaffolding in some environments, especially workplaces, can perpetrate a kind of 'mind invasion' when it sculpts behaviour and habit in ways that harm the the interests of employees. Liao and Huebner (2020) have argued that material artefacts and environments can themselves be oppressive, when biased in the same direction as an oppressive system, and bi-directionally causally embedded it. We differ from Slaby in considering both cognitive and affective scaffolding. We also differ from Slaby, and from Liao and Huebner, as we will see, because hostility is a narrower and less systemic notion than oppression. We say more about how hostile scaffolding relates to mind invasion and oppressive things, and how our proposal complements theirs, in section (5).

2. Scaffolding, benign and hostile

2.1 Benign Scaffolding

The general idea of cognitive scaffolding is of something *external* — usually to the body, but sometimes to the brain — that in some way supports cognitive processes. As Clark put it in *Being There,* "exploitation of external structure is what I mean by the term *scaffolding*." (Clark 1997, p45). There's generally little appetite for regimenting scaffolding talk, which is recognised to pick out a varied collection of situated, embodied and distributed cognitive phenomena. Given our aim of making a significant addition to the range of things recognised as scaffolding, we need to be more deliberate. Here is an inclusive and neutral characterisation: *Cognitive Scaffolding* is external structure that changes the cognitive demands of a task.

To count as *external* something should be to a significant extent outside the 'skull or skin'. *Structure* is an inclusive place-holder, embracing objects, mechanisms, temporally organised processes, symbolic systems, public language, and the activity of other possibly encultured agents. We also intend *cognitive* to be understood inclusively, to embrace affective and motivational phenomena as well as more traditionally cognitive ones, and to to include cases with a 'pushmi-pullyu' character where the cognitive and motivational or affective are entangled (Millikan 1995). That said, we will sometimes emphasise whether scaffolding is predominantly affective or cognitive. Because the cognitive thus understood is a large and varied category, many kinds of *change* can be brought about by scaffolding. A jacket left hanging on the bathroom door handle scaffolds remembering to dress smartly in the morning. A cook organising partly prepared ingredients in a workspace reduces search time by simplifying perception, or facilitates the construction of a pleasing regular arrangement with reduced measurement and planning (Kirsch 1995). Hutchins (1995) describes many examples of artefacts which transform the computational demands of navigation, including the 'three scale nomogram' (example in Hutchins 1995, p148), which converts several multiplication and division operations relating distance time and speed into drawing a straight line that connects two known values and which crosses all of three appropriately positioned logarithmic scales. Slide rules on which several rigid logarithmic scales — some mobile in relation to others — are marked allow multiplication and division to be performed by manipulating the scales, because the physical structure of the device preserves a large number of mathematical relationships (Hutchins 1995, p170f).

Although the above examples emphasise the traditionally cognitive (memory, perception, reasoning), scaffolding can also be affective. Consider how mementos, places, pieces of recorded music and forms of embodied activity can help occasion or sustain desired affective or motivational states, and sometimes have downstream cognitive effects (Griffiths & Scarantino 2009; Colombetti & Kruger 2015; Piredda 2019; Maiese 2016). Finally, saying that scaffolding *changes* the cognitive demands of a task doesn't presume that the task could be performed unaided. We emphasise that one consequence of this characterisation is that whether or not something is scaffolding depends on and is restricted to the specified task and agent, and isn't an all or nothing property of the external structure. This is not a bug, but a feature: scaffolding, like the affordances of ecological psychology, isn't purely environmental, but rather irreducibly relational. The slide rule is only useful to those competent to use it, and one person's motivating exercise playlist might be another's acoustic purgatory.

The idea of scaffolding borders on and overlaps with the notions of epistemic actions, and cognitive niches. Epistemic — as distinct from pragmatic — actions are "physical actions that make computation easier, faster, or more reliable" (Kirsh & Maglio 1994, 513-4). A common example is manually rotating a jigsaw puzzle piece so that determining whether it fits is a relatively easy visual matching task, compared to mental rotation. Some scaffolding is produced by epistemic actions (leaving the jacket on the bathroom door, arranging the workspace), but this isn't a general requirement for scaffolding which can be found as well as made. A cognitive niche, on the other hand, is a significantly scaffolded environment that may benefit many individuals. This version of the more general idea of niche construction and inheritance emphasises the ways in which the activity of some living things makes their environments more cognitively favourable for themselves and their descendants. Beaver dams, for example, increase the length of the water/land boundary, and so increase the foraging space close to the safety of water (Sterelny 2011, p239). Neither epistemic actions nor scaffolding are uniquely human. Sterelny offers a hawk that "chooses a roost which maximises its view of its hunting territory" and the laying down of pheromone trails between next and food sources by some ant species as examples of epistemic action (Sterelny, 2010, p470). The former case exploits the scaffolding potential of a perch favourable to one individual and its task, while in the latter combined activity of many individuals contributes to making a cognitive niche or scaffolded space in which simple pheromone-cued dispositions can efficiently gather food. The trails 'carry information about direction and distance' and, given the trail making and following dispositions, the intensity of a trail 'carries information about the value of the food resource' (Sterelny 2003, p19, drawing on Hölldobler & Wilson 1990, pp272-3). Our focus here is on scaffolding rather than either epistemic actions or niches, even though some scaffolding might be produced by or depend on epistemic actions, and even though an accumulation of scaffolding might add up to a niche.

We called the characterisation above *neutral* to signal that it doesn't require that scaffolding be helpful to the agent engaged in the task. That scaffolding is helpful generally goes without saying, and examples offered in the literature to defend the claim that cognition is scaffolded, are overwhelmingly helpful. We call such scaffolding benign, and fill this out as follows: *Benign Cognitive Scaffolding* is external structure that changes the cognitive demands of a task in ways that serve the interests of the agent attempting the task.

Just as we count something as scaffolding at all in a way that is task and agent relative, whether it is benign is interest relative. Scaffolding is benign when it serves the agent's *actual* interests, which needn't be known to or appreciated by the agent. When those interests can be clearly identified, and are themselves consistent, then whether the scaffolding is benign or not can be correspondingly clear. It often is clear, as the examples reviewed above, and the wider literature on scaffolding, shows. It may also sometimes be difficult to determine, for example when a task contributes to multiple goals in heterogenous ways, or there are competing credible accounts of the agent's interests. A slide rule, for example, might augment my computational powers, while unpleasantly reminding me of a cruel mathematics teacher many years ago. We propose two strategies for dealing with challenging cases. The more difficult one is to defend a position on the agent's *overall* interests in order to show how they are being benefited by the scaffolding given the task. The simpler alternative is to take the success conditions of the task itself to determine a *local* interest, which scaffolding either serves or does not. We'll use both strategies in what follows.

Many examples of benign scaffolding involve cues or surface indicators in an environment: objects placed as reminders, pheromone trails, road markings and conventional signage for bathrooms and elevators in public spaces. While using some of these cues might depend on significant on-board resources, including competence with symbolic systems or public language, considered *as scaffolding* they are relatively superficial. Not all scaffolding is like this - the slide rule that converts mathematical operations into physical manipulations, also requiring some numerical competence from the user, isn't merely a favourably situated label or cue, but the instantiation of a large number of structured relationships that works *because* the normal range of manipulations preserve those relationships. That is, subject to the markings being correctly read, the manipulations can do significant computational work. When it matters to draw the distinction, which is very much one of degree, we'll call the former, superficial scaffolding 'shallow' and the instances where significant computational work gets done externally, including transforming inferences of one kind into manipulations or processes of a different kind, 'deep'. The deep shallow distinction is independent of the identification of scaffolding as benign, or — as we'll see shortly — hostile.

Examples of deep benign scaffolding are readily found. Consider a month planner with days of the week arranged in regular columns or rows, with already planned activities filled in. The same set of facts could be stored on scraps of paper — one for each engagement — and kept in a bag marked 'April'. By mapping calendar time onto spatial relationships in an orderly way, the month planner transforms tedious, time-consuming and error-prone enquiries into quick visual searches and inferences. (Compare using the bag of notes and the planner to identify two adjacent, available weekdays suitable for scheduling an event.) Navigation maps distort some aspects of physical space in order to preserve others, for example so that straight lines on the map pick out locations with shared directional relationships, at the cost of accurately representing scale (Hutchins 1995, Chapter 3). Many underground train maps distort distances and directions so as to facilitate queries and inferences focused on topological relationships between passenger stations and different services (Kent 2021). That using these resources typically requires some skill from the user doesn't make a difference to the key point: desk planners, slide rules, some maps, numeral systems and routines for long-division, etc., are forms of deep scaffolding which depend on the structure of the scaffolding itself, and whether it successfully preserves (including under manipulation) a version of some favoured set of relationships. This is a difference of degree, and allows boundary cases. An individual spot of ant pheromone is a plausible instance of shallow scaffolding, but an accumulation of them into a *trail* can carry additional information because of their relationships with each other and further facts. A pheromone trail is less shallow than an individual spot.

The claim that there can be deep scaffolding is a version of the claim that cognition itself can be extended, a thesis often defended with functionalist considerations. Functionalists hold that mental or cognitive states are distinguished at least primarily by the causal and functional relations they stand in with inputs, behaviours and other cognitive states. What Wheeler (2010) calls the 'chauvinism-busting' property of multiple realizability holds that systems that stand in the right kinds of causal and functional role, even though they might work or be composed in importantly different ways, can be cognitively equivalent. The label *extended* functionalism makes explicit that one of the chauvinisms being busted is internalism about cognition. The extended functionalist does not require or presuppose that the functional chains implementing cognition remain bounded by skull or skin, and is, for example, willing to grant that features of the anatomy and position of one or more legs in conjunction with local sensors can constitute parts of the loops in a neural network controller while taking a different form to the internal 'wired' connections (Beer et al 1993, discussed in Clark 1997).

2.2 Hostility and Scaffolding

We can now characterise hostile scaffolding. This isn't merely scaffolding that fails to be benign by being inefficient or unreliable, and so doesn't serve the interests of the agent attempting a task. A calendar that somehow encouraged inaccurate inferences, or consumed more time than it saved, could fall short of being benign, but wouldn't thereby serve another agent. Hostile scaffolding both fails to be benign in undermining the interests of the agent attempting the task *and* by doing so serving the interest of another agent: *Hostile Cognitive Scaffolding* is external structure that changes the cognitive demands of a task in ways that undermine the interests of the agent attempting the task, and serve those of another agent.

We take the term hostile here, as noted, from Sterelny (2003) who distinguished between informational environments as follows. In an *informationally transparent* environment signals an organism can detect are reliably good occasions for behaviours it is capable of producing, so cue-driven behaviour will be successful (Sterelny 2003, p. 20). Environments aren't reliably transparent. When relevant features of the environment "map in complex, one to many ways onto the cues [an organism] can detect" they are *informationally translucent* (Sterelny 2003, p. 21). Sometimes the translucency is not the result of indifferent heterogeneity in the world, but is produced by other living things with competing interests, in which case the environment is *informationally hostile*. That is, hostility is not merely a matter of antagonistic or competing interests existing, but concerns their expression through information. (We're only taking Sterelny's account of hostility from this, and won't make further reference to the notion of transparency.)

Sterelny's emphasis on hostility is a corrective to over-optimistic slogans in artificial intelligence and robotics about letting the world be its own model, or own best representation (Brooks 1991; Clark 1997, p46). Control systems based on responding to cues depend on the reliability of the cues, but their predictable responsiveness is an opportunity for exploitation. Sterelny surveys examples including female *Photuris* fireflies that engage in aggressive mimicry by producing the mating signals of females of other firefly species in order to attract, kill and eat males of those species (Sterelny 2003, p15; Lloyd 1997). He argues that when the investment pays its way expanded cognitive capacities that are less cue-bound, including proto-beliefs and proto-desires, can be a strategy for dealing with translucence and hostility. It is also central to his view that humans have long been distinctive in the extent of their reliance on scaffolding, and in their cultivation and maintenance of cognitive niches that support mastering the extraordinary informational demands of their lifestyles (Sterelny 2010, 2012). Given his emphases on hostility and scaffolding, it might seem surprising that Sterelny doesn't develop an account of hostile scaffolding. Rather, he's argued it is unlikely to be a serious problem, focusing on how shared resources are made reliable by being shared, and that the costs of exploiting by means of scaffolding are unlikely to be covered by the returns (Sterelny, 2010, p474). We return to his reasons for scepticism, and respond to them, in section (4.3) below.

Just as whether scaffolding is benign is relative to task, agent and interests, so whether it is hostile is relative to the task and interests of one agent, and the interests of another. Recall the slide rule that extends my calculating capacities while bringing painful memories. If my task is being in a good mood, then the ruler might fail to be benign (it certainly isn't helping) but that doesn't make it hostile, because nobody is benefiting. Hostile scaffolding, if there is any, has both victims and beneficiaries. Since only the victim need be engaged in a task, determining whether the scaffolding benefits another agent will depend on the of the beneficiary and their interests. In section (3) below we look at this in cases which fix interests by reference to gene replication, and in section (4) we link interests to returns in money.

3. Extended Cognition and Extended Phenotypes

Dawkins (1982) argues that *phenotypes* reach beyond the skin. This claim applies to construction such as spider webs, and to the *behaviour of other organisms*. We argue that the extended phenotype thesis and extended functionalism about cognition (the latter including the possibility of deep scaffolding and extended cognition) are complementary in hitherto neglected ways, and together predict the *possibility* of deep hostile scaffolding. The complementarity is twofold: First, extended functionalism provides additional ways of thinking about manipulation and behaviour control than standard expositions of the extended phenotype thesis. Second, the extended phenotype thesis pays close attention to contested control and attempted manipulation, which is salutary for standard expositions of extended cognition focusing on benign extensions.

The extended phenotype thesis is a development of the gene-centric perspective on natural selection, which prioritises the interests of replicators (in being replicated) over those of the organism in which they occur (Dawkins 1976a). This perspective is defended against individual-level selectionism by theoretical considerations, such as that only relatively digital gene sequences have enough copying fidelity and potential 'immortality' to be the target of selection, and by its explaining cases where genes replicate at the expense of the interests of their host organism.[3] The gene-centric perspective coincides with the individual organism one, when the respective interests align, so claims about extended phenotypes can often be formulated independently of the gene-centric perspective. Nothing that follows here depends on the accepting the gene-centric view.

The 'phenotype' of an organism is commonly understood as the *bodily* expression, in morphology, physiology and behaviour, of the genotype. Conventional thinking was that phenotypes were bounded by the skin. As with extended cognition, the extended phenotype thesis is defended by arguments that common assumptions about boundaries are incorrect, and that we will be unable to explain important phenomena unless we ditch these assumptions. In one case, the presumption is that cognition is bounded by skin or skull, in the other it is that the phenotype is bounded by the skin — or scales, bark, etc. — of the organism. Both deploy the extended functionalist thought that standing in appropriate causal relationships is more important than falling on one or the other side of some boundary (Dawkins 1976b, p8). In the case of gene-centric selection, the relevant causal relationships are effects (under typical genetic, developmental and environmental conditions) on the rate at which a gene is replicated. So, Dawkins points out, some organisms reliably modify the world outside their own skins in ways that are crucial to their biological success, such as beavers building dams. Their capacities and dispositions to make these modifications, including wood-cutting teeth and disliking the sound of running water, can have a heritable (genetic) basis. When they do the constructions should be considered as part of the phenotype irrespective of whether they are 'inside' the host's body or not. So, if there replicators that under typical conditions increase the likelihood of building dams, or building them this way rather than that, then there are 'genes for' dam building, and dams are part of the phenotype. Dawkins argues that causal relationships should guide our thinking, whether or not the chains of causation cross the skin: "…an animal artefact, like any other phenotypic product whose variation is influenced by a gene, can be regarded as a phenotypic tool by which that gene could potentially lever itself into the next generation." (1982, p199). Whether the levering is

3 One example of this is the 'Medea' gene (for *Maternal effect dominant embryonic arrest*) in flour beetles, which produces – in effect – both a toxin and its antidote. An individual with the gene suffers no net effect, besides paying the cost of producing a molecule it goes on to deactivate, and those without it are normal. A heterozygous female beetle expresses *Medea* (producing the toxin) in her germline, where only half of her young produce the antidote, while the others die during larval development (Beeman et al 1992). This phenotypic difference does *not* serve the reproductive interests of heterozygous individuals.

achieved by 'adorning the tail of a male bird of paradise with a sexually attractive blue feather' or by 'causing a male bower bird to paint his bower with pigment crushed in his bill out of blue berries' doesn't matter. Levers are multiply realised, and can be extended.

Some extended phenotypes involve extended cognition or benign scaffolding. Dawkins urges us to see a spider's web as "a temporary functional extension of her body, a huge extension of the effective catchment area of her predatory organs" (1982, p198). Clark, when defending extended cognition, repeatedly draws an analogy with Dawkins' use of this example (e.g. 1997, 2008, p118; pp241-242). Webs and dams are constructions, one category of extended phenotype. The other involves manipulating the behaviour of other organisms to serve replicators in another. Here an example, not from Dawkins. The *Lymantria dispar* virus causes infected European gypsy moth larvae to climb up their host trees before dying, liquefying and releasing spores of the virus. The behaviour of uninfected individuals includes reducing predation risk by hiding in tree bark, or climbing down to the soil (Hoover et al 2011). Dawkins discusses various real and hypothetical examples, including the increased tendency of those infected by rabies to bite, which would plausibly serve transmission of the pathogen which is carried in saliva, but does no good for the host (Dawkins 1982, p220). He offers as the 'central theorem' of his view that: "An animal's behaviour tends to maximize the survival of the genes 'for' that behaviour, whether or not those genes happen to be in the body of the particular animal performing it." (Dawkins 1982, p233)

Dawkins' account of genetic 'action at a distance' focuses on *how* a gene in one organism could influence the behaviour of another organism. Sometimes we can be confident influencing is happening, including with cuckoo eggs and cuckoo chicks which depend on successfully getting past the defences of nesting birds subject to brood parasitism. Feeding a cuckoo chick does not advance their interests at all, but is very much in the interests of the cuckoo. Dawkins considers two mechanisms (sometimes speculatively illustrated) to manipulate behaviour: sensory cues, and neural interventions like drugs, or short-circuits. In the cuckoo case, he emphasises cues: The parasitised reed-warbler "is an active, complex machine, with sense organs, muscles and a brain. The brood parasite must […] infiltrate the defences of the host's nervous system, and its ports of entry are the host's sense organs." (Dawkins 1982, p68)

Cuckoo brood parasitism evolved independently three times (Payne 2005) and the various arms races between host and parasite species differ widely. Cuckoo chicks eject other eggs and chicks, but still need hosts to feed them more than they would one of their own (smaller) chicks. Some of this is achieved through exaggerated versions of species-specific feeding cues. Rufous bush chats in Spain, where cuckoo chicks have orange gapes, bring more food to chicks of their own whose mouths have been dyed orange (Alvarez 2004). Orange dye makes no difference to reed-warbler parents in England, who instead bring more food to their own young in response to recordings of cuckoo chick begging calls (Davies, Kilner & Noble 1998; Davies 2015, Chapter 10). These are exemplary cases of informational hostility, exploiting cue-sensitive behavioural dispositions (see Sterelny 2003, p27-28).

Sensory cues aren't the only ways to exert influence because, says Dawkins, a "nervous system can be subverted if treated in the right way" (1982, p69) including being "vulnerable to manipulation by a clever-enough pharmacologist" (1982, p71). The scenarios Dawkins entertains for genetic 'action at a distance' include influence by pheromones and by parasites within host nervous systems, and all suppose that the downstream effects of replicators must somehow get at the *brain* or *nerves* of the organism to influence it, whether they do so through the sensory front door or a neural or metabolic service elevator. The founding treatment of extended phenotypes, that is, does not contemplate extended cognition. Japyassú and Laland (2017) go so far as to suggest that Dawkins' internalism about cognition is analogous to his gene-centrism. Whatever the reason, we maintain that extended functionalism about cognition complements the extended phenotype view: If cognition can be extended and distributed beyond the skin, including by means of deep scaffolding, others means of action at a distance may be available. The possibility that interests us is influence by hostile scaffolding, that is external structure that transforms the cognitive demands of a task in a way serving another agent.

Notice that a case of an extended phenotype operating through hostile scaffolding would have to satisfy two conditions. First, features of the hostile scaffolding would have to be among the downstream effects of a gene. Second, the interests to be served are the replication of instances of that gene, through further downstream effects of the exploited agent interacting with the scaffolding. The functional chain has to go from a gene through the scaffolding and back to the gene. These are important restrictions from the gene-centric perspective, which has a standard way of handling all questions about interests. But neither is required for hostile scaffolding more generally. First, our fairly permissive notion of scaffolding (section 2 above) allows for scaffolding that is both found and made. Road markings are scaffolding for all road users, not merely for those who paint them. Second, we're willing to contemplate any reasonably defended position about the actual interests served, or to associate interests with the success conditions of the task being attempted. There could consequently be examples of hostile scaffolding that don't count as extended phenotypes.

The second restriction above, making a difference to the replicators responsible for the extended phenotypic difference does important work in Dawkins' account, because without something like it all of the varied downstream effects of genes would be counted as phenotype (Dawkins 1982, pp. 233-4). A similar danger threatens our characterisation of hostile scaffolding, because of the possibility that arbitrary agents would gain in some way from merely inefficient scaffolding that failed to serve the interests of its user. Suppose, for example, that a person is so captivated by an appealing time-management tool to scaffold their scheduling that they take time away from the very work they were trying to organise, and are beaten by a competitor. The scaffolding fails to serve their interests *and* another agent benefits. This consequence is unwelcome, because the idea of hostility that we are taking from Sterelny (2003) concerns the expression of competition in an informational environment. We therefore require that the agent benefitting be 'in the loop' that influences the placing or significant features of the structure that is a candidate for hostile scaffolding. So, if the competitor had recommended the time-management tool hoping for it to be counter-productive, we'd have a case of attempted hostile scaffolding, and if successful a real example.

We earlier distinguished between shallow and deep scaffolding, where the latter involved significant processing happening outside the skin. This allows us to distinguish less and more interesting versions of the claim that there can be hostile scaffolding, including among non-human animals. The less interesting version concerns shallow scaffolding. Cases where the activity of an organism modifies its own environment to support behaviour selection, as some insects do with pheromone trails, count as benign scaffolding or cognitive niche-construction. Examples of manipulation where other organisms produce relevantly similar cues in ways that bias activity to serve the interests of the manipulators would then count as hostile scaffolding. Both cases transform the cognitive environment and influence activity but differ in whose interests are served. (This is what the 'central theorem' of the extended phenotype warns us not to take for granted.) It is less interesting to call manipulation by cues hostile scaffolding, because that's merely renaming what is amply covered by existing accounts of aggressive mimicry, cryptic colouration and other phenomena, covered in Sterelny's (2003) account of the threats of hostility to cue-bound behaviour, and by Dawkins' own treatment of manipulation by cues (1982). This part of our view may be salutary for the over-optimistic literature on human scaffolding, but only delivers one of the two complementarities between extended phenotype and extended cognition that we identified.

Both come into play with deep scaffolding, where significant computational or processing work depends on external structure. The non-obvious and more interesting claim is that *deep hostile scaffolding* is possible. This would go beyond the 'sensory cues or drugs' range of options Dawkins offers for thinking about action at a distance, and require extended functionalism about cognition. We admit not having plausible non-human examples of deep hostile scaffolding. Hypothetically, a creature that laid malicious pheromone *trails* to lead ants to their death would constitute deep hostile scaffolding, if we count the benign trails as deep scaffolding (section 2.1 above). In the following section we argue that there are real *human* examples of deep hostile scaffolding.

4. Real hostile scaffolding

Our main examples of deep hostile scaffolding are drawn from electronic gambling machines and aspects of casino management. The gambling customer might be pursuing multiple goals, including recreation or entertainment, and cannot safely be presumed to have the goal of winning money overall. Most of our discussion therefore focuses on a specific task that arises periodically for the gambler: determining whether to continue wagering. This might be answered deliberately or by pre-reflective or implicit appraisal - specifying the task doesn't commit to a method of resolution. The decision whether to not to continue, though, is a point where the interests of player and house can come apart, because the player who stops gambling stops losing to the house.

Rachlin wrote that in "a way (not to be taken too literally) the casino serves the decision theorist, as the heavens have served the physicist, as a natural laboratory." (1989, p105-6). His point was that games of chance played for money provided a series of choice episodes with definite timing and magnitudes of cost and payout in a single modality, offering clean data for the decision scientist analogous to how the mostly frictionless motion of approximately spherical celestial bodies served physics. Casinos and gambling machine designers, however, don't merely *study* choice, they aim to exploit it. Some of the tools deployed to this end involve ambience, layout, lighting and sound, and a leading handbook of casino management (Friedman 2000) favours thinking of casinos as mazes (in Schüll 2012, p39). Much of this effort is aimed at prolonging gambling episodes. Some of the means involve scaffolding, some of it predominantly affective and motivational, more more conventionally cognitive.

4.1 Affective Hostility

Affective scaffolding that is unreliably benign isn't hard to find. Technology in smart watches and other devices that track activity or proxies for stress level, for example, *could* contribute to an entirely beneficial improvement in someone's self-care. When these tools involve significant external processing, fed back to the user, they can count as deep scaffolding. One reason to doubt that such gadgets are reliably benign is provided by Nguyen (2020, Chapter 9). He argues that technologies that 'gamify' some areas of life have insidiously distorting effects on the agency of a user seduced by the simplicity of the numerical feedback into neglect of the more demanding and less certain practical reasoning that rational agency requires. If he's right, the scaffolding could fail to be benign for some tasks, but fall short of making it hostile if the harms were by-products without beneficiary. The same tracking technology could become clearly hostile if, hypothetically, used by an exploitative employer to increase the efficiency of value extraction in ways harmful to the interests of employees. Slaby (2016) argues for something like this, although he's primarily concerned with the downstream effects of what we would call shallow (affective) hostile scaffolding.

Casino player tracking technology is a credible example of deep hostile scaffolding (Schüll 2012, Chapter 5). The first player tracking system in casinos was implemented in 1985 and mimicked 'loyalty' systems from airlines and banks (Schüll 2012, p144). Players were given punch cards that were notched with every jackpot. After enough notches, cards could be redeemed various rewards, providing an incentive to keep gambling, and gathering data for casinos. Contemporary tracking is much more sophisticated. Casinos now track the value of every bet, a player's win and lose rate, the rate of button pressing on machine games, the timing of breaks, and details of food and drink purchases, over a player's entire history, perhaps in multiple venues (Schüll 2012, p144). In 2005, Harrah formulated a plan to "optimize" player value with a system estimating, given historical data, how much a player could lose over how long while still continuing play. The system predicted a 'pain point' for the individual player, and dispatched a 'Luck Ambassador' to give out vouchers (for further gambling, meals, or show tickets) shortly before the pain point (Schüll 2012, p154). That system backfired because players, especially the most valuable ones, didn't like being interrupted (Schüll 2012, p169f), but less obtrusive and disruptive interventions — including directly and instantly distributing non-cashable credits to the players at individual machines, and allowing refreshments to be ordered from the machines — have taken their place. Some machines disable animations or other features when players play fast enough, adjusting their behaviour to the revealed preference for a rate of interaction (Schüll 2012, p168-9), and multiline machines offer frequent small 'wins' (some of them net losses) in ways that encourage prolonged play (Dixon et al 2014). Colombetti and Krueger (2015) include gambling machines and environments in their examples of affective scaffolding, and note how these machines, especially ones responsive to player behaviour 'induce an extreme state of absorption and isolation' in players, and are aware that this is in the service of casino revenues. Their main focus is on arguing that affectivity too, is sometimes scaffolded, and they use machine gambling environments as a powerful example of entrenchment into a corporal schema. We want to emphasise that the scaffolding can also be hostile, and the induced state of absorption harmful or associated with increased vulnerability. Some gambling researchers, seeking to highlight this, have called the flow-like state induced in some machine gamblers 'dark flow' (Dixon et al, 2017).

The stated aim, and common effect, of most of these innovation is to prolong gambling, and reduce the interval between gambling sessions. Revenue from gambling machines, which provide the most detailed player tracking data, is a function of 'time on device' and rate of play, which are what venues and designers seek to maximise. With the increasing sophistication of the machines themselves, machine gambling has sharply risen in the fraction of casino revenue it contributes, the fraction of casino space allocated to it, and the number of available machines. At the same time, machine gambling addiction has become more common, and come to account for the majority of cases of gambling addiction, with the onset of addiction in cases of machine gambling being dramatically quicker than for other forms of gaming (Schüll 2012, p4-6, 14-21).

Player tracking systems coupled with attempts to predict the cessation of play, and postpone it with interventions that cost the house less than it expects to extract, are exemplary cases of deep hostile scaffolding. This is, as noted above, a conflict played out over the task of deciding whether to continue gambling, where player interests and those of the house come apart. It is 'scaffolding' because the influence on player behaviour is facilitated through external structures that transform task demands. It is 'deep' in the sense that casinos are gathering and processing historical and real-time player data to estimate the affective and motivational states of players and manipulate upcoming events (such as those shortly before the predicted pain point) to prolong play. It is 'hostile' because the main goal is extracting more money by prolonging play and not serving the player's own interests.

4.2 Cognitive Hostility

The hostility in contemporary gambling technology is not limited to the affective and motivational, and includes compelling instances of deep hostile cognitive scaffolding. To make this case we turn to selected features of the development of slot machines, leaving aside poker-based and other games some of which involve different forms of hostile scaffolding. The earliest slot machines were mechanical. Several adjacent spinning cylindrical reels with visible symbols at intervals, and a number of stopping points for each reel, were set up so that they would pay out under certain conditions. The 1899 'Liberty Bell', for example, had three reels each with five symbols, and one stopping point for and between each symbol. This machine accepted and paid out in nickels after the reels were set in motion by a player pulling a handle. (One nickel to play, ten dispensed when all three reels stopped on bells.) The reels were stopped by a braking system and timing bar, so each stopped individually. The small number of possible outcomes meant that modest payouts were required to keep the machine profitable (Schüll 2012, p80).

Early mechanical slot machines had some benign scaffolding-like features. The number of symbols on the reels, their spacing and motion, could be observed through the viewing window, and reliably related – in fair machines, with repeated observation – to the actual odds. Apparent near misses, for example where only one reel was one stop away from a payout configuration, were genuine. Later electronic gambling machines broke the connection between the visible reels and the odds, allowing it to be distorted, and for apparent near-misses to multiply as the odds of winning went down. One key innovation was using electronic random number generators to determine the outcome of play. This gave designers and owners more precise control over the odds than mechanical systems, a change partly sold as a way of blocking tampering. After the introduction of random number generators, the visible spinning was no longer the process that culminated with the outcome of play, instead it was an animation reverse-engineered to present the already processed output of the random number generator.

This made possible the innovation called 'virtual reel mapping' (Schüll 2012, p81f; Enkvist 2009, p166f; Telnaes 1984). This brings it about that the odds of winning are different from, and can be considerably *lower* than, what the appearance of the reels suggests. Lower actual odds enable larger, more tempting jackpots to be offered while keeping machines profitable. It was trivial to make the actual odds low by having the random number generator select an outcome from as large a range as the machine designer or casino wanted. Virtual reel mapping was a solution to the problem of relating the output of the random number generator configured to give lower odds of winning to the activity of the simulated reels, while keeping the appearance of the machine the same. The key is that the randomization process, over a larger set of outcomes than is actually displayed, is 'mapped' onto the simulated reels in a biased way, with disproportionately more of the virtual stops mapped onto "low-paying or non-paying blank positions" on the visible reel (Schüll, 2012, p87). The larger the number of non-paying stops, the more non-paying combinations, and the larger the jackpot that could be offered.

The intent to separate the motion and stopping of the visible reels from the actual odds is explicit in the patent, which introduces the innovation as "enabling any odds to be set without changing the physical characteristics of the machine" (Telnaes 1984, p1), and explains that "the purpose of this invention [is] to increase the capability of the designer to include high payoffs without increased physical size of the machine and with uniform presentation of the games of different models to the player" (Telnaes 1984, p8). It was possible to lengthen the odds with an 'honest' display by having more or larger reels. But players *correctly* perceive such machines "as being less "good" in terms of winning and payout chances" (Telnaes 1984, p8). So, said Telnaes, "it is important to make a machine that is perceived to present *greater chances of payoff that it actually has* within the legal limitations that games of chance must operate" (1984, p8, emphasis added). The visible reels retain the scaffolding-like properties found in mechanical machines. The movement of apparently rigid cylinders with regularly spaced symbols encourage inferences about chances, and contribute to what the patent calls player 'acceptance'. But the inferences encouraged are inaccurate because the actual odds differ from what the reels suggest, and the deal players are encouraged to 'accept' by the appearance of the machine is considerably more skewed against their interests (in expected money) than it seems. This is deep hostile scaffolding, leading to prolonged play.

A good question at this point is why this works so well, given that prolonged play is also an opportunity to learn from experience of losing to the house. Part, by no means all, of the answer in some machines is the extent of investment in encouraging trust in the scaffolding features themselves. Slot machine manufacturers have paid considerable attention to cultivating the sense of interaction with a genuine physical apparatus. The manufacturer Bally once held a leading position because their machines' handles could be pulled more or less quickly, with apparent haptic feedback about the progressive loading of reel springs (Schüll, 2012, p83; Enkvist 2009, p164f). The virtual reel mapping technology posed challenges for the gambling machine industry because it broke that link, and companies explored different ways of *simulating* convincing mechanical interaction, and filed various patents aimed at achieving the feel of the mechanical handle, without giving players any genuine control. Ekvist quotes a registered patent that referred to the "certain amount of feel and *presumed control* over the device" (Ekvist 2009, p165, emphasis added). Handles were eventually found to slow down play unacceptably, and button-push play initiation came to dominate. Other attention focused on the movement of the reels controlled by step-motors, or the animation of simulated reels, again to encourage the sense of genuine physical interaction. One solution was to layer semi-transparent computer graphics over genuine but blank spinning reels, another to stack liquid crystal displays to enhance the appearance of depth (Schüll 2012, p83). Yet another allowed players to stop the apparent spinning. Although this made no difference to the outcome, "gamblers using such "stop" features seem to feel they have an effect on outcomes and […] to persist at play for significantly longer periods" (Schüll 2012, p84).[4] This sense of interacting with something that behaves like a genuine physical object, even when it isn't one and when its actual dynamics are different and concealed, is important to the coupling (Sutton 2010), or the complementarity (Menary 2010), that makes these machines operate as deep scaffolding even when serving the operator over the player.

4 One patent that includes attention to controlling the system when players have the option to stop reels, observes that it "is generally known in the art that the movement of the symbols up to four frames after the actuation of the stop button does not give an unnatural impression to the player", and describes a strategy to control displayed outcomes consistent with avoiding this 'unnatural impression' (Okada 1986, p27).

The fact that the sometimes simulated reels are controlled *after* outcome determination, allows another form of misrepresentation. We noted above that on a fair mechanical machine, apparent near misses are genuine. Near misses are also motivating (a point suggested in Skinner 1953, p397), leading to extended machine playing time (Strickland & Grote 1967, Reid 1986, Clark et al 2009)[5] and may be distinctively motivating for disordered gamblers (Sescousse et al 2016). One of the earliest innovations in mechanical slots was expanding the viewing window vertically, allowing more genuine near misses to be seen (Schüll 2012, p80). When the reel mapping and display are entirely under designer control apparent near misses can be made more frequent. This requires stopping the reels so that the final configuration looks as though a small difference – at minimum a single reel stopping one step earlier or later – would have led to a winning play. One Nevada industry dispute centred on this, because a line of machines which presented apparent near-misses several times more often than chance, by yet another patented method, were gaining market share from competitors (Harrigan 2008, Enkvist 2009).[6] Although the outcome was that some kinds of fabricated near miss were ruled unacceptable, this doesn't affect our point about deep hostile scaffolding, which is that an apparatus designed to cultivate a strong sense of having the dynamics of a real physical machine, so as to encourage visual properties being taken as proxies for probabilities, was used to exploit players to the advantage of casinos. In addition, the 'banning' of fabricated near-misses is restricted, and misleadingly frequent apparent near-misses an inevitable by-product of virtual reel mapping. This is because the grouping of several non-paying virtual stops either side of a winning symbol will routinely make that symbol appear to stop 'one stop away' from a winning position more frequently (Harrigan 2008). The 'legal limitations' at work in this industry fall far short of preventing many forms of deception and manipulation.

That most users of slot machines play voluntarily, and that the activity is presented and often understood as a form of entertainment, does not affect our point here. To reiterate, hostility need not be an all or nothing affair, and is relative to a task and the interests of two agents. The customer who enters the casino for fun, willing spend some time and lose some money while having it, still faces the periodic task of determining whether they have had enough yet. And the house gains if that question doesn't arise at all, or is answered negatively.

---

5 That there is a 'near miss' effect is not universally accepted. See Pisklak, Yong & Spetch (2020).

6 Universal Distributors patented a system which generated high numbers of apparent near misses, and received regulatory approval for it. "Universal's "near-miss" software increased the excitement during play and extended the average length of each gambling session, therefore resulting in a higher player appeal" (Enkvist 2009, p174, Okada 1986).

Suppose we consider the overall interests of customers, instead of focusing narrowly on the task of deciding whether to continue wagering at any point. While gambling addicts might have a conflicted relationship to their gambling, it is uncontroversial that someone who isn't gambling addict has no interest in becoming one. The considerations we noted above about the faster onset and increased prevalence of gambling addiction with the rise of the new machines suggest that the overall interests of customers are not being served. Hostile scaffolding helps turn recreation into exploitation.

Our position here complements an account offered by Ross (2020) who draws a useful contrast between non-human mammal and human encounters with targets of addiction. Elephants and baboons can and do get drunk when they find low-toxicity sources of alcohol, but "they are at no risk of addiction […] because they cannot cultivate sources of low-toxicity alcohol. Their parties are windfalls, the frequency of which they cannot influence" (Ross 2020, p. 6). Technological innovation has enabled humans, Ross contends, to 'engineer addictive environments', including by processing and stockpiling alcohol, nicotine and other substances, and by building environments, including casinos, that foster addiction. As Ross points out, part of how this works in the machine gambling case is because our reward learning systems find the cycling of action and feedback offered by gambling highly reinforcing, but are unable to 'settle on a model of genuine randomness' (Ross 2020, p. 3). He argues that it's not reasonable to say that there is anything *wrong* with mammalian learning systems, including those of gambling addicts, given that evolutionary history did not require them to cope with anything like electronic gambling machines. Rather, we should recognise the role of exploitative environments, including the exploitative learning environments that are the 'business model' (Ross 2020, p. 7) of casinos, in explaining addiction. We add only that casinos aren't merely exploitative learning environments, but are augmented with deep hostile scaffolding.

4.3 An objection

We noted above that Sterelny, although sensitive to the importance of both scaffolding and hostility, has said that he doesn't think manipulation of public epistemic resources is likely. He does say that "hostile manipulation of [our] informational environment is a serious danger", however sees the danger as restricted it to a limited set of interactions, mostly comprising "one-on-one high-stakes negotiations" (Sterelny, 2010, p474), such as someone exploiting or harming the notebook-using Otto from Clark and Chalmers (1998) by erasing or altering Otto's external memory. But for Sterelny, manipulation involving public resources (such as deceptively changing the maps in a subway station) is unlikely, because sharing itself increases reliability, and the fact that many agents use different copies of the resources at unpredictable times makes it difficult to exploit a chosen target. "In many circumstances, public domain resources cannot safely be used to manipulate a specific target for a specific purpose" (Sterelny, 2010, p474). We'll take these remarks as an objection, giving reasons to think that hostile scaffolding isn't *likely*.

Sterelny is of course correct that much scaffolding is public, and characterised by considerable redundancy. One of his examples is maps of underground train systems, which are duplicated in stations, train carriages and printed media. Those resources scaffold the environment of people faced with the task of planning travel, and the maps themselves are often examples of deep scaffolding that organise information in ways that simplify a selected class of inferences (see section (2) above). It would take substantial effort and cost to manipulate such systems in order to misdirect a single passenger, and is consequently unlikely to bring sufficient returns. Another point that he makes, referring to the Monty Python sketch in which someone maliciously produces a Hungarian-English phrasebook leading users to make inappropriate and ridiculous utterances ('My hovercraft is full of eels'), and suggests that some of the humour here arises because "it is difficult to envisage circumstances in which an author would gain from producing a maliciously misleading phrasebook, for an author cannot know when, where, by whom or with what effect such a book will be read" (Sterelny 2010, p275). Here, too, Sterelny is correct: Even if an environmental intervention will mislead the agent who happens to rely on it, that doesn't mean that the perpetrator will benefit.

The considerations that Sterelny emphasises here aren't, however, fully general. They limit the reach of hostile scaffolding, but do not exclude it entirely. In the case of affective scaffolding in casinos, a key part of the story is that the casino itself has extraordinarily comprehensive control over the whole environment, including its layout and design, and what opportunities for various kinds of activities are to be found within it. Casino operators have conducted extensive research on ambience and spatial arrangements (Schüll 2012, chapter 1), and the player-tracking tools that seek to extend 'time on device' monitor players within these highly controlled environments. Players enjoy none of the protections of redundancy, but must muddle along in a hostile niche. The casino's target isn't a specific individual that needs to be picked out of a crowd but anyone and everyone who enters, *and* they have the capacity to track individuals in detail within their venues. Not only that, achieving similar levels of control over an environment doesn't require command over a large physical space (as with casino-managers) because an increasing number of the relevant environments people face today *virtual*, and the control has only to be achieved within a social media app, operating system or computer game. The potential victims of a variety of forms of exploitation carry increasingly powerful and permanently connected computing devices around with them. So unlike the deceptive train map scenario Sterelny imagines, the targets broadcast their location and behaviour, and carry the instruments of their possible bespoke exploitation around with them. In many of these digital environments the very tools to optimise time on device developed in gambling have been deployed or modified, leading Harris, former Google design ethicist, to call a smartphone 'The Slot Machine in Your Pocket' (Harris 2016). That users often don't directly pay in money is a dis-analogy with the casino case, but is irrelevant to the hostility, because much connected media and games extract advertising advertising revenue that is a function of time on device. People on social media applications also have determine when to stop, and when to resume, and in both cases their interests and those of providers can come apart. The malicious phrasebook prank is 'ballistic', its imagined perpetrator isolated from the consequences for victims. The casino knows who it is manipulating, as do social media and search companies who also much of their users' individual history, and what devices they are holding. To state the point here more generally, hostile scaffolding will pay its way when there are ways of placing it around potential targets, and the costs are more than covered by the returns. The considerations that fix the boundaries between profitable and unprofitable scaffolding, and the challenges of matching scaffolding to targets, are not fixed, but vary with available technology. Increasingly powerful, cheap, ubiquitous and connected computing devices make the robustness of train maps small comfort indeed. Sterelny's general warning to take hostility seriously stands, and optimism about scaffolding being reliably benign is in no better shape than optimism that the world is its own best representation.

5. Conclusion

Clark, who greatly raised the profile of scaffolding, once wrote that the "single most important task, it seems to me, is to better understand the range and variety of types of cognitive scaffolding, and the different ways in which non-biological scaffoldings can augment (or impair) performance on a task" (Clark 2002, p29). The parenthetical 'or impair' is a rare interruption to his usually optimistic focus on benign scaffolding, and is repeated in the same paragraph when he says that the "Holy Grail here is a taxonomy of different types of external prop, and a systematic understanding of how they help (and hinder) human performance" (Clark 2002, p29; also Clark 2010, p58f). We take our treatment of hostile scaffolding to be a contribution to the task Clark highlighted, focused on his passing hints about unhelpful scaffolding, and significantly inspired by Sterelny's (2003) working out of the importance of informational hostility.

We observed near the start of this paper that our claims are not entirely unprecedented, and we close by noting what we view as complementarities and differences with some valuable recent work. Liao and Huebner (2020), first, have argued persuasively that there are 'oppressive things', that is material artefacts and environments which contribute to oppression, when their effects are biased in the same direction as, and they are causally embedded in, an oppressive system. Liao and Huebner endorse two of the four components of 4E cognition - that cognition is *embedded*, and *embodied* - say that they are 'sympathetic' to a third - that it is *enactive* - but say that neither of them is 'particularly committed' to the fourth - that cognition is *extended* (2020, n3). This isn't necessarily a rejection of the possibility of deep scaffolding, because they're only distancing themselves from 'first wave' extended mind thinking (see Sutton 2010), but that is our main focus. Their concern with oppression, understood as a relationship between groups, is different from our understanding of hostility which relates to individual agents. Liao and Huebner consequently don't count an imagined biased slot machine as oppressive because it would harm 'all players equally' (2020, p8), although they are also aware that the organisation of gambling industries exploits and maintains oppressive class relations. Our respective accounts will therefore sometimes pick out different examples, with hostile scaffolding including environmental structure that exploits individuals indifferently to their group membership. But they may also complement each other. In cases where something that works like hostile scaffolding makes more sense in relation to the interests of groups than individuals, it may be better to refer to oppressive scaffolding.

Slaby's (2016) account of 'mind invasion' is a further partial precedent. He argues that affective scaffolding in some environments, especially workplaces, can amount to a kind of 'mind invasion' when it sculpts behaviour and norms in ways that are detrimental to the interests of employees. Slaby criticises much thinking about extended cognition (and extended minds) for taking for granted a 'user/resource model' which presumes that external structure will serve the interests of the agents encountering it. This line of criticism is somewhat analogous to Dawkins' argument against assuming that the interests served by behaviour are those of the agent performing the behaviour itself, and our own argument for taking hostile scaffolding seriously. Like Liao and Huebner, Slaby's intention is partly political - in Slaby's case drawing on Protevi's reading of Deleuze (Protevi 2012, p132). Again, this means that 'mind invasion' and hostile scaffolding may pick out different examples, with the latter including cases of individual exploitation beyond workplaces. In addition Slaby's account focuses on affective scaffolding that we'd count as shallow, even if its accumulated downstream effects can be substantial. So as well as picking out different examples, the case we've developed there covers different ground in considering scaffolding that is cognitively hostile, as well as deep in the sense of doing significant cognitive work.

In summary, while the literature on situated and extended cognition is dominated by examples of cognitive scaffolding that are *benign,* in changing the cognitive demands of a task in ways that favour the agent performing the task, we should not assume that this is generally true. We've argued that scaffolding can also be *hostile*, in serving the interests of another agent. The notion of deep hostile scaffolding potentially fills a gap in extended phenotype thinking about the possible routes to behaviour manipulation, and - when freed from the gene-centric account of interests - suggests routes to manipulation and exploitation that go far beyond populating the environment with misleading superficial cues. Finally, we've provided extended examples from casinos and some types of electronic gambling machines to argues that there are genuine examples of deep hostile scaffolding. If we're right, further work is called for, including work estimating how much hostile scaffolding there is, who is gaining and losing, what varied forms it takes, and what defences we may have or be able to develop.

References

Alvarez, F. (2004) The conspicuous gape of the nestling Common Cuckoo *Cuculus canorus* as a supernormal stimulus for Rufous Bush Chat *Cercotrichas galactotes* hosts. *Ardea*, 92: 63-68.

Beeman, R.W., Friesen K,S., & Denell, R.E. (1992) Maternal-Effect Selfish Genes in Flour Beetles. *Science,* 256(5053): 89–92.

Beer, R., Chiel, H., Quinn, K., Espenschied, S., & Larsson, P. (1992) A distributed neural network architecture for hexapod robot locomotion. *Neural Computation* 4(3): 356- 365.

Brooks, R.A. (1991) Intelligence without representation, *Artificial Intelligence*, 47: 139–159.

Cazden, C. (1981). Performance before competence: Assistance to child discourse in the zone of proximal development. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 3, 5–8.

Clark, A. (1997) *Being There: Putting Mind, Brain and Body Together Again*. Cambridge, Massachusetts: MIT Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 7–19.

Clark, A. (2002) Towards a science of the bio-technological mind. *International Journal of Cognitive Technology*, 1(1): 21–33.

Clark, A. (2008) *Supersizing the Mind*. Oxford University Press.

Clark, A. (2010) Memento's revenge: The exended mind extended. In R. Menary (Ed.), *The extended mind*. Cambridge: MIT.

Clark, L., Lawrence, A. J., Astley-Jones, F., & Gray, N. (2009). Gambling near-misses enhance motivation to gamble and recruit win-related brain circuitry. *Neuron*, 61(3), 481-490.

Colombetti, G. & Kruger, J. (2015) Scaffoldings of the affective mind. *Philosophical Psychology*, 28(8), 1157-1176.

Davies, N.B., Kilner, R.M., & Noble, D.G. (1998) Nestling cuckoos, Cuculus canorus, exploit hosts with begging calls that mimic a brood. *Proceedings of the Royal Society B,* 265: 673-678.

Davies, N. (2015) *Cuckoo: Cheating by Nature*. Bloomsbury.

Dawkins, R. (1976a) *The Selfish Gene*. Oxford University Press.

Dawkins, R. (1976b) Hierarchical organisation: a candidate principle for ethology. In

Growing Points in Ethology (eds P. P. G. Bateson & R. A. Hinde) Cambridge University Press, pp. 7-54.

Dawkins, R. (1982) *The Extended Phenotype*, Oxford: Oxford University Press.

Dixon, M.J., Graydon, C., Harrigan, K.A., Wojtowicz, L., Siu, V., & Fugelsang, J.A. (2014) The allure of multi-line games in modern slot machines. *Addiction*, 109(11), pp. 1920–1928.

Dixon, M.J., Stange, M., Larche, C.J., Graydon, C., Fuglesang, J.A., & Harrigan, K.A. (2017) Dark Flow, Depression and Multiline Slot Machines. *Journal of Gambling Studies*, 34, pp. 73-84.

Enkvist, M. (2009) *Creating Player Appeal: Management of Technological Innovation and Changing Pattern of Industrial Leadership in the U.S. Gaming Machine Manufacturing Industry, 1965– 2005*. PhD diss., Department of Economic History, School of Business, Economics and Law, University of Gothenburg.

Friedman, B. (2000) *Designing Casinos to Dominate the Competition*. Reno, NV: Institute for the Study of Gambling and Commercial Gaming.

Griffiths, P., & Scarantino, A. (2009). Emotions in the wild: The situated perspective on emotion. In M. Aydede & P. Robbins (Eds.), The Cambridge handbook of situated cognition (pp. 437–453). Cambridge: Cambridge University Press.

Harrigan, K.A. (2008) Slot Machine Structural Characteristics: Creating Near Misses Using High Symbol Award Ratios. *International Journal of Mental Health and Addiction* 6: 353– 68.

Harris, T. (2016) The Slot Machine in Your Pocket, *Spiegel Online* (July 27, 2016, 5:25pm), https://www.spiegel.de/international/zeitgeist/smartphone-addiction-is-part-of-the-design-a-1104237.html

Hölldobler, B. & Wilson. E. (1990) *The Ants*. Harvard University Press.

Hoover, K., Grove, M., Gardner, M., Hughes, D.P., McNeil, J., & Slavicek, J. (2011) A Gene for an Extended Phenotype, *Science*, 333(6048): 1401.

Hutchins, E. (1995) *Cognition in the Wild*. Cambridge, Massachusetts: MIT Press.

Japyassú, H.F., & Laland, K.N. (2017) Extended Spider Cognition. *Animal Cognition* 20(3): 375–95.

Kent, A. J. (2021) When Topology Trumped Topography: Celebrating 90 Years of Beck's Underground Map. *The Cartographic Journal*, 58(1), pp. 1-12.

Kirsh, D. & Maglio, P. (1994) On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*, 18: 513-549.

Liao, S. & Huebner, B. (2000) Oppressive Things, *Philosophy & Phenomenological Research*. DOI: 10.1111/phpr.12701

Lloyd, J. E. (1997) Firefly Mating Ecology, Selection and Evolution. In J.C. Chloe and B.J. Crespi (Eds) *Mating Systems in Insects and Arachnids*. Cambridge: Cambridge University Press.

Maiese, M. (2016) Affective Scaffolds, Expressive Arts, and Cognition. *Frontiers in Psychology*, 7:359.

Menary, R. (2010) The extended mind and cognitive integration. In R. Menary (Ed.), *The extended mind*. Cambridge: MIT.

Millikan, R. (1995) Pushmi-pullyu representations. *Philosophical Perspectives, 9: AI, Connectionism and Philosophical Psychology*, pp. 185-200.

Nguyen, T. (2020) *Games: Agency as Art*. Oxford University Press.

Okada, K. (1986) Slot machine with random number generation. U.S. patent 4573681

Payne, R. B. (2005) *The Cuckoos*. Oxford University Press.

Piredda, G. (2019) What is an affective artefact? A further development in situated affectivity. *Phenomenology and the Cognitive Sciences*. [Online first] https://doi.org/10.1007/s11097-019-09628-3

Pisklak, J.M., Yong, J.J.H., & Spetch, M.L. (2020) The Near-Miss Effect in Slot Machines: A Review and Experimental Analysis Over Half a Century Later. *Journal of Gambling Studies*, 36: 611-632.

Protevi, J. (2013) *Life War Earth: Deleuze and the Sciences*. Minneapolis, MN: University of Minnesota Press.

Rachlin, H. (1989) *Judgement, Decision, and Choice*. New York: W.H. Freeman and Company.

Reid, R. L. (1986) The psychology of the near miss. *Journal of Gambling Behavior*, 2(1), 32–39.

Ross, D. (2020) Addiction is socially engineered exploitation of natural biological vulnerability. *Behavioural Brain Research, 386*. https://doi.org/10.1016/j.bbr.2020.112598

Schüll, N.D. (2012) *Addiction by Design: Machine Gambling in Las Vegas*. Princeton, NJ: Princeton University Press.

Sescousse, G., Janssen, L.K., Hashemi, M.M., Timmer, M.H.M., Geurts, D.E.M., ter Huurne, N.P., Clark, L., & Cools, R. (2016) Amplified striatal responses to near-miss outcomes in pathological gamblers. *Neuropsychopharmacology,* 41: 2614-23. doi:10.1038/npp.2016.43

Skinner, B. F. (1953) *Science and human behavior*. New York: Macmillan.

Slaby, J. (2016) Mind Invasion: Situated Affectivity and the Corporate Life Hack. Frontiers in Psychology 7:266. *doi:10.3389/fpsyg.2016.00266*

Sterelny, K. (2003) *Thought in a Hostile World*, Oxford: Blackwell.

Sterelny, K. (2010) Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences*, 9(4): 465-481.

Sterelny, K. (2011) Externalism, Epistemic Artefacts and The Extended Mind. In Richard Shantz (Ed.) *The Externalist Challenge*. Berlin: De Gruyter.

Sterelny, K. (2012) *The Evolved Apprentice*. MIT.

Strickland, L. H., & Grote, F. W. (1967). Temporal presentation of winning symbols and slot machine playing. *Journal of Experimental Psychology*, 74, 10–13.

Sutton, J. (2010) Exograms, interdisciplinarity and the cognitive life of things. In R. Menary (Ed.), *The extended mind*. Cambridge: MIT.

Telnaes, I.S. (1984) Electronic gaming device utilising a random number generator for selecting reel stop positions. US Patent 4448419.

Vygotsky, L. S. (1978). Mind and Society. Cambridge, MA: Harvard University Press.

Wheeler, M. (2010) In Defense of Extended Functionalism. In R. Menary (Ed.), *The extended mind*. Cambridge: MIT.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100.